On Computational Power of Quantum Branching Programs

Farid Ablayev* Aida Gainutdinova† Marek Karpinski‡

Abstract

In this paper we study a model of Quantum Branching Program (QBP) and investigate its computational power. We define several natural restrictions of a general QBP model, such as a read-once and a read-k-times QBP, noting that *obliviousness* is inherent in a quantum nature of such programs.

In particular we show that any Boolean function can be computed deterministically (*exactly*) by a read-once QBP in width $O(2^n)$, contrary to the analogous situation for quantum finite automata. Further we display certain symmetric Boolean function which is computable by a read-once QBP with $O(\log n)$ width, which requires a width $\Omega(n)$ on any deterministic read-once BP and (classical) randomized read-once BP with permanent transitions at each level.

We present a general lower bound for the width of read-once QBPs, showing that the upper bound for a considered symmetric function is almost tight.

1 Introduction

Richard Feynman observed in 1982 ([9]) that certain quantum mechanical effects cannot be simulated effectively on a classical computer. This observation led to a general idea that perhaps computation based on quantum

*Max-Planck Institute for Mathematics, Bonn. Visiting from University of Kazan. Supported in part by Russia Fund for Basic Research 99-01-00163 and Fund "Russia Universities" 04.01.52. ablayev@ksu.ru

†Dept. of Computer Science, University of Kazan. Supported in part by Russia Fund for Basic Research 99-01-00163 and Fund "Russia Universities" 04.01.52.aida@ksu.ru

‡Dept. of Computer Science, University of Bonn, and Mathematical Sciences Research Institute, Berkeley. Supported in part by DFG grants, DIMACS, IST grant 14036 (RAND-APX), and Max-Planck Research Prize. Research at MSRI was supported by NSF grant DMS-9810361. marek@cs.uni-bonn.de

effects could be much more efficient than the classic one. During the last decade the area of research of quantum computation was an intensively growing area. Shor's quantum algorithm for factoring integers [15] that runs in polynomial time is well known now.

As mentioned in [4] quantum computers may consist of two parts: a quantum part and a classical part with communication between them. In such a case, a quantum part could be considerably more expensive than the classical part. Therefore, it might be useful to construct a quantum part as simple as possible. This motivates a study of restricted models of quantum computation.

During the last decade many different restricted quantum computation models have been investigated. In particular the quantum analogs of Boolean circuitsNC and ACC [13, 8] has been introduced. Another model — quantum finite automata has been introduced and first studied by Kondacs and Watrous [12], see [7] for more information on the subject. It has been shown that uniform one-way quantum finite automata with bounded error probability cannot accept all regular languages [12]. But Ambainis and Freivalds [4] presented a regular language which can be computed by a quantum finite automaton with bounded error probability of exponentially smaller size than the corresponding classical randomized finite automaton. One of the more recent papers comparing classical and restricted quantum computation models is [16].

A classical branching program (BP)(see, e.g., [17]) is a convenient nonuniform model for studying various restricted variants of computational models. Leveled oblivious permutation BPs are well known in the complexity theory, their computational power is remarkable (deterministic leveled oblivious permutation BPs with constant width have the same power as $\log n$ depth circuits [6]). It seems also that the branching programs are well suited for comparing restricted quantum models with their classical counterparts.

Nakanishi, Hamaguchi, and Kashiwabara introduced also in [14] a variant of quantum model of BPs as an extension of probabilistic BPs. In this paper we introduce a model of quantum BPs different from that of [14]. Without loss of generality we consider leveled BPs. For a leveled BP P, all the paths in a BP are of the same length, and one can move only from the nodes of the i-th level to the nodes of the $i+1$-th level. We denote by $w(P)$ a width of P. That is, $w(P)$ is a maximum number of nodes on various levels of P. In our case a superposition of P is any element $|\psi\rangle = \sum_{i=1}^{w(P)} z_i|q_i\rangle$. We notice that we may need much less quantum bits than in a model of [14].

In ths paper we investigate a restricted computational variant of a quantum branching program— a quantum read-once BP. First we show that a read-once (exact) quantum BPs (noting that the *obliviousness* is inherent in a quantum nature of such programs) can compute an arbitrary Boolean function. Next we display a certain symmetric Boolean function which is computable by a read-once QBP with $O(\log n)$ width, which requires a width $\Omega(n)$ on any deterministic read-once BP and on any (classical) randomized read-once BP with permanent transitions at each level. We present a general lower bound for the width of read-once QBPs, showing that the upper bound for a considered symmetric function is almost tight.

2 Preliminaries and Definitions

We consider a d-dimensional Hilbert complex space $\mathcal{H}^d$ with a norm $||.||$. That is, for $z \in \mathcal{H}^d$, $z = \{z_1, \ldots, z_d\}$, $||z|| = \sqrt{\sum_{i=1}^{d} |z_i|^2}$.

Recall some basic notations from the quantum mechanics. A pure quantum state (or superposition) of a quantum system QS with d stable states $\{1, \ldots, d\}$ (d-state QS) can be expressed by associating an amplitude z_i (a complex number) to each state i of a QS. Quantum mechanics uses for that the following notations. Consider the quantum *basis states* $\{|1\rangle, \ldots, |d\rangle\}$ where $\{|i\rangle\}$ is the set of d-dimensional orthogonal basis vectors of $\mathcal{H}^d$ where $|i\rangle$ denotes the unit vector with the value 1 at i and 0 elsewhere. A pure quantum state or a *configuration* of a QS can be written as

$$|\psi\rangle = \sum_{i=1}^{d} z_i |i\rangle$$

or just $|\psi\rangle = (z_1, \ldots, z_d)$ where $|\psi\rangle \in \mathcal{H}^d$. The specific notation $|\psi\rangle$ corresponds to the Dirac 'ket'-notation for a column-vector $(z_1, \ldots, z_d)$. An element z_i of $|\psi\rangle$ is called an amplitude of the basis state $|i\rangle$ of a QS, and $|z_i|^2$ is the probability of finding a QS in the state i when QS is beeing measured.

A time evolution of configurations of a QS in discrete steps is reversible and conveniently expressed using Heisenberg's matrix mechanics. That is, if in a current step a configuration of a QS is $|\psi\rangle$, then in the next step a configuration of a QS would be $|\psi'\rangle$ where $|\psi'\rangle = U|\psi\rangle$ and U is a $d \times d$ unitary matrix.

We are going to define now a *quantum transformation* as follows. Let $X = \{x_1, \ldots, x_n\}$ be a set of Boolean variables. Define quantum transformation (d-dimensional quantum transformation) on $|\psi\rangle \in \mathcal{H}^d$ as a triple

$\langle j, U(0), U(1)\rangle$ where j is the index of a variable $x_j \in X$, and $U(0), U(1)$ are reversible transformations of $\mathcal{H}^d$ presented by unitary $d \times d$ matrices. A quantum transformation $\langle j, U(0), U(1)\rangle$ of $|\psi\rangle$ acts as follows: $U|\psi\rangle = |\psi'\rangle$. If $x_j = 1$ then $U = U(1)$ else $U = U(0)$.

2.1 Definition of a QBP

A Quantum Branching Program of width d and length l ((d, l)-QBP) based on a QS is defined by a triple

$$P = \langle T, |\psi_0\rangle, F\rangle$$

where T is a sequence (of length l) of d-dimensional quantum transformations of a d-state QS:

$$T = (\langle j_i, U_i(0), U_i(1)\rangle)_{i=1}^{l},$$

with $|\psi(0)\rangle$ an initial configuration of P, and $F \subseteq \{1, \ldots, d\}$ a set of accepting states.

We define a computation of P over an input $\sigma = \sigma_1, \ldots, \sigma_n \in \{0, 1\}^n$ as follows:

1. Computation of P starts from the superposition $|\psi_0\rangle$. On the i-th step, $1 \le i \le l$, P transforms a superposition $|\psi\rangle$ to a superposition $|\psi'\rangle = U_i(\sigma_{j_i})|\psi\rangle$.
2. After the l-th (last) step of a quantum transformation, P measures its configuration $|\psi_\sigma\rangle$ where $|\psi_\sigma\rangle = U_l(\sigma_{i_l})U_{l-1}(\sigma_{i_{l-1}}) \ldots U_1(\sigma_{i_1})|\psi_0\rangle$. The measurement is presented by a diagonal zero-one projection matrix M where $M_{ii} = 1$ if $i \in F$ and $M_{ii} = 0$ if $i \notin F$. The probability $p_{accept}(\sigma)$ of P accepting an input σ is defined by

$$p_{accept}(\sigma) = ||M|\psi_\sigma\rangle||^2.$$

We call a QBP P read-once, if each variable $x \in \{x_1, \ldots, x_n\}$ occurs in a sequence T of quantum transformations of P at most once.

2.2 Computation of a Boolean Function

- A QBP P is said to compute (with an unbounded error) a Boolean function $f_n : \{0, 1\}^n \to \{0, 1\}$ if for all $\sigma \in f^{-1}(1)$ the probability of P accepting σ is greater than $1/2$ and for all $\sigma \in f^{-1}(0)$ the probability of P accepting σ is at most $1/2$.

- A QBP P computes a Boolean function f_n with a bounded error if there exists an $\varepsilon > 0$ such that for all $\sigma \in f^{-1}(1)$ the probability of P accepting σ is at least $1/2 + \varepsilon$ and for all $\sigma \in f^{-1}(0)$ the probability of P accepting σ is at most $1/2 - \varepsilon$. We call ε a margin, and say that P $(1/2 + \varepsilon)$-computes f_n.
- We say that a QBP P exactly computes f_n if P computes f_n with the margin $1/2$ (with the zero error probability).

3 Computational Properties

The following property of a simulation of QBPs with complex valued amplitudes by QBPs with real valued amplitudes is similar to simulations by quantum Turing machines cf. [5].

Property 1 *Let a* (d, l)-QBP P *with complex valued amplitudes computes (with an unbouded error, with a bounded error, exactly) a function* f_n. *Then, there exists a* $(2d, l)$-QBP P' *with real amplitudes within the interval* $[-1, +1]$ *that computes (with an unbouded error, with a bounded error, exactly, respectively) the same function* f_n.

Proof: Let us consider a product $U|\psi\rangle$ of a complex valued $d \times d$ matrix U

$$U = \begin{pmatrix} z_{1,1} & \dots & z_{1,d} \\ \vdots & & \vdots \\ z_{d,1} & \dots & z_{d,d} \end{pmatrix}$$

where $z_{i,j} = a_{i,j} + \sqrt{-1}b_{i,j}$,
with a d-dimensional complex valued vector $|\psi\rangle$

$$|\psi\rangle = \begin{pmatrix} z_1 \\ \vdots \\ z_d \end{pmatrix}$$

where $z_i = a + \sqrt{-1}b_i$.
The above product can be simulated by a product of a $A|v\rangle$ of $2d \times 2d$ real valued matrix A

$$A = \begin{pmatrix} Z_{1,1} & \dots & Z_{1,d} \\ \vdots & & \vdots \\ Z_{d,1} & \dots & Z_{d,d} \end{pmatrix}$$

where $Z_{i,j} = \begin{pmatrix} a_{i,j} & b_{i,j} \\ -b_{i,j} & a_{i,j} \end{pmatrix}$ is matrix presentation of the complex number $z_{i,j}$, and
$|v\rangle$ is a $2d$-dimensional real valued vector

$$|v\rangle = \begin{pmatrix} a_1 \\ b_1 \\ \vdots \\ a_d \\ b_d \end{pmatrix}$$

We notice that all amplitudes z in superpositions $|\psi\rangle$ of P are such that the real a and imaginary b part of $z = a + \sqrt{-1}b$ are from intervals $[-1, +1]$. From the above it is easy to conclude that a d-dimensional l-length complex valued quantum transformations of a d-state QS

$$(\langle j_i, U_i(0), U_i(1)\rangle)_{i=1}^{l},$$

can be simulated by the corresponding $2d$-dimensional l-length real valued quantum transformations of a $2d$-state QS

$$(\langle j_i, A_i(0), A_i(1)\rangle)_{i=1}^{l}.$$

□

Below we show that a bounded error read-once QBPs are powerful enough to compute an arbitrary Boolean function. By contrast, we notice thati the uniform one-way quantum finite automata when accepting with a bounded error probability can compute only a proper subset of regular sets [12]. See also [7] for more recent results on the complexity properties of quantum finite automata.

Property 2 *For arbitrary Boolean function f_n, there exists a read-once $(2^n, n)$-QBP that exactly computes f_n.*

Proof: The proof is evident. The following read-once $(2^n, n) -$ QBP P satisfies our proposition. All possible configurations of P are trivial. That is, a configuration $|\psi\rangle$ of P contains exactly one 1, and all the rest components of $|\psi\rangle$ are 0. The initial configuration of P is $|\psi_0\rangle = (1, 0, \ldots, 0)$. P reads input variables in order $x_1, x_2, \ldots, x_n$.

In each step $i, 1 \leq i \leq n$, P reads input σ_i and transforms its current configuration $|\psi\rangle$ as follows. If $\sigma_i = 0$ then $|\psi\rangle$ does not change. If $\sigma_i = 1$, then the 1 of the configuration $|\psi\rangle$ is "moved" to 2^{n-i} positions to the right in the next configuration $|\psi'\rangle$.

For an input sequence $\sigma = \sigma_1, \dots .\sigma_n$, denote by l_σ the number of position of 1 in the final (after reading σ) configuration of P. Clearly we have that $l_\sigma \neq l_{\sigma'}$ iff $\sigma \neq \sigma'$.

Now determine the set of accepting states F of P as follows: if $f(\sigma) = 1$, then $q_{l_\sigma} \in F$. If $f(\sigma) = 0$, then $q_{l_\sigma} \notin F$. □

Denote by **EP**-QBP$_{const}$ the class of all Boolean functions exactly computed by the constant width and polynomial length (in the number of function variables) QBPs.

Property 3 *For the complexity class* $\mathbf{NC^1}$ *it holds that*

$$\mathbf{NC^1} \subseteq \mathbf{EP}\text{-QBP}_{const}.$$

Proof: Proof is evident by a known result of Barrington [6]. Having a permutation deterministic branching program P of width 5 computing a Boolean function f_n it is easy to construct a $(const, poly)$-QBP P' which exactly computes f_n. □

Consider now the following symmetric Boolean function MOD_{p_n}: For an input $\sigma = \sigma_1, \dots, \sigma_n \in \{0,1\}^n$ we have $MOD_{p_n}(\sigma) = 1$ iff a number of ones in σ is divisible by p_n, where p_n is a prime and $p_n \leq n/2$.

Theorem 1 *The function MOD_{p_n} can be computed by a read-once $(O(\log p_n), n)$-QBP with a one-sided error probability.*

The proof of this theorem will be presented in the section below. We have clearly that any deterministic OBDD for MOD_{p_n} has $\Omega(p_n)$ width.

3.1 Proof of Theorem 1

The proof is similar to that of [4]. [4] introduces the following regular set L_P. For a prime p, a language L_p over a single letter alphabet is defined by $L_p = \{u : |u| \text{ is divisible by } p\}$. It is proved that for any $\varepsilon > 0$, there is a QFA with $O(\log p)$ states recognizing L_p with probability $1 - \varepsilon$.

We construct a QBP P accepting the inputs $\sigma \in MOD_{p_n}^{-1}(1)$ with the probability 1 and rejecting the inputs $\sigma \in MOD_{p_n}^{-1}(0)$ with the probability at least 1/8. Consider a $(2,n)$-1QBP P^k for $k \in \{1,\dots,p_n-1\}$. A quantum program $P^k = \langle T^k, |\psi_0^k\rangle, F^k\rangle$ is based on the following 2-state quantum system, $T^k = (\langle i, U^k(0), U^k(1)\rangle)_{i=1}^n$ where

$$U^k(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, U^k(1) = \begin{pmatrix} \cos(2\pi k/p_n) & -\sin(2\pi k/p_n) \\ \sin(2\pi k/p_n) & \cos(2\pi k/p_n) \end{pmatrix}, |\psi_0^k\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, F^k = \{1\}.$$

Denote by $l(\sigma)$ a number of 1-s in the sequence σ, $l(\sigma) = \sum_{i=1}^n \sigma_i$.

Lemma 1 ([4]) *After reading an input $\sigma = \sigma_1, \dots, \sigma_n$, the superposition of P^k is*

$$|\psi\rangle = \cos\left(\frac{2\pi\, l(\sigma)k}{p_n}\right)|1\rangle + \sin\left(\frac{2\pi\, l(\sigma)k}{p_n}\right)|2\rangle.$$

Proof: The proof follows from the description of a QBP P^k.

□

If the number of ones in an input σ is divisible by p_n, then $2\pi\, l(\sigma)k/p_n$ is a multiple of 2π and $\cos\Big(2\pi\, l(\sigma)k/p_n\Big) = 1$, $\sin\Big(2\pi\, l(\sigma)k/p_n\Big) = 0$. Therefore all QBPs P^k accept inputs $\sigma \in f_{n,p_n}^{-1}(1)$ with probability 1.

Following [4], we call P^k "good" for input $\sigma \in MOD_{p_n}^{-1}(0)$ if P^k rejects σ with probability at least 1/2.

Lemma 2 *For any $\sigma \in MOD_{p_n}^{-1}(0)$, at least $(p_n-1)/2$ of all P^k are "good".*

Proof: According to Lemma 1, after reading an input $\sigma = \sigma_1, \dots, \sigma_n$ the superposition of P^k is

$$|\psi\rangle = \cos\left(\frac{2\pi\, l(\sigma)k}{p_n}\right)|1\rangle + \sin\left(\frac{2\pi\, l(\sigma)k}{p_n}\right)|2\rangle.$$

Therefore, the probability of accepting the input $\sigma \in MOD_{p_n}^{-1}(0)$ is $\cos^2\Big(2\pi\, l(\sigma)k/p_n\Big)$. $\cos^2\Big(2\pi\, l(\sigma)k/p_n\Big) \le 1/2$ iff $\Big|\cos\Big(2\pi\, l(\sigma)k/p_n\Big)\Big| \le 1/\sqrt{2}$. This happens if and only if $\Big(2\pi\, l(\sigma)k/p_n\Big)$ is in $[\pi/4 + 2\pi j, 3\pi/4 + 2\pi j]$ or in $[5\pi/4 + 2\pi j, 7\pi/4 + 2\pi j]$ for some $j \in N$. $\Big(2\pi\, l(\sigma)k/p_n\Big) \in [\pi/4 + 2\pi j, 3\pi/4 + 2\pi j]$ iff $\Big(2\pi\,(l(\sigma)k \bmod p_n)/p_n\Big) \in [\pi/4, 3\pi/4]$. p_n is prime, and $l(\sigma)$ is relatively prime with p_n. Therefore, $l(\sigma) \bmod p_n$, $2l(\sigma) \bmod p_n, \dots$, $(p_n-1)l(\sigma) \bmod p_n$ is $1, 2, \dots, p_n-1$ in different order. Consequently, it is enough to find the power of a set $I = \{i_1, \ldots, i_l\} \subset \{1, \ldots, p_n - 1\}$ such that $2\pi\, i_j\, / p_n \in [\pi/4, 3\pi/4]$ or $2\pi\, i_j\, / p_n \in [5\pi/4, 7\pi/4]$. Since p_n points $2\pi/p_n, \ldots, 2\pi(p_n - 1)/p_n, 2\pi$ are regularly distributed on a circumference, and sectors $[\pi/4, 3\pi/4]$ and $[5\pi/4, 7\pi/4]$ are exactly a half of the circumference, we have $|I| \geq \lfloor p_n/2 \rfloor \geq (p_n - 1)/2$.

□

Following [4], we call a set of quantum programs $S = \{P^{i_1}, \ldots, P^{i_t}\}$ "good" for $\sigma \in f^{-1}_{n,p_n}(0)$ if at least 1/4 of all its elements are "good" for this σ.

Lemma 3 *There is a set S of 1QBPs with $|S| = t = \lceil 16 \ln p_n \rceil$ which is "good" for all inputs $\sigma \in MOD^{-1}_{p_n}(0)$.*

Proof: We consider the following procedure $\mathcal{A}$ for a construction of a set S.

> For a fixed input σ with $l(\sigma) \leq p_n - 1$, $\mathcal{A}$ selects a quantum branching program uniformly at random from $\{P^1, \ldots, P^{p_n - 1}\}$.

The probability of selecting a "good" QBP at each step is at least 1/2. Using Chernoff inequality, we have that the probability that less than 1/4 fraction of all QBPs from the set S are "good", for any fixed σ with $l(\sigma) \leq p_n - 1$, is at most

$$\exp((-16 \ln p_n)/2(1/2)^2/2 = 1/p_n.$$

Hence the probability that a constructed set is not "good" for at least one input σ with $l(\sigma) \leq p_n - 1$, is at most $(p_n - 1)/p_n > 0$. Therefore there exists a set which is "good" for all inputs σ with $l(\sigma) \leq p_n - 1$. This set is "good" for the inputs σ with $l(\sigma) > p_n$ as well, since any QBP P^k returns a current superposition of a starting superposition after reading all p_n ones, and hence it works the same way on inputs σ, σ' with $l(\sigma) = l(\sigma') \bmod p_n$.

□

We construct a 1QBP P accepting inputs $\sigma \in MOD^{-1}_{p_n}(1)$ with probability 1 and rejecting inputs $\sigma \in MOD^{-1}_{p_n}(0)$ with probability at least 1/8 as follows. A 1QBP P consists of QBPs from "good" set $S = \{P^{i_1}, \ldots, P^{i_t}\}$, which work in parallel. In the starting superposition of P all these programs have equal amplitudes.

The inputs $\sigma \in MOD_{p_n}^{-1}(1)$ are always accepted with the probability 1 because all the P^ks accept them. For any input $\sigma \in MOD_{p_n}^{-1}(0)$ at least 1/4 of all $P^k \in S$ reject it with probability at least 1/2, and the total probability of rejecting any $\sigma \in MOD_{p_n}^{-1}(0)$ is at least 1/8.

We can make the error as small as possible using a standard technique for reducing an error probability for a one-sided error computation. That is, we take $d = d(\varepsilon)$ copies of such a 1QBP P and run them uniformly at random. In this case the width of 1QBP will be $O(\log p_n)$. □

- We call a branching program P *stable* if its transformations do not depend on the level of P.

From the proof of our theorem we have that the constructed QBP for MOD_{p_n} is a stable branching program.

Corollary 1 *The function MOD_{p_n} can be computed by a stable read-once $(O(\log p_n), n)$-QBP with a one-sided error.*

Below we show that MOD_{p_n} function is hard, in fact, for the randomized OBDDs.

3.2 Lower Bound for Randomized OBDDs for MOD

Randomized OBDDs were introduced and firstly investigated in [1], see also [17].

Theorem 2 *Any stable probabilistic OBDD computing MOD_{p_n} has width at least p_n.*

Proof: Assume that there is a stable probabilistic OBDD P of width $q < p_n$ computing MOD_{p_n} with probability $1/2 + \varepsilon$ for a fixed $\varepsilon \in (0, 1/2]$. We can assume without loss of generality that each level of P has exactly q nodes. Let $\mu^j = (\mu_1^j, \ldots, \mu_q^j)$ be a probability distribution of states of P on the j-th level, where μ_i^j is the probability of being in the i-th node of the j-th level. We can describe a computation of P on an input $\tilde{\sigma} = \sigma_1, \ldots, \sigma_n$ as follows:

- A computation of P starts from the initial probability distributions vector μ^0.

- At the j-th step, $1 \le j \le n$, P reads an input σ_{i_j} and transforms a vector μ^{j-1} to $\mu^j = \mu^{j-1}A$, where A is a $q \times q$ stochastic matrix, $A = A(0)$ if $\sigma_{i_j} = 0$ and $A = A(1)$ if $\sigma_{j_k} = 1$.

- After the last (n-th) step of the computation, P accepts the input $\tilde{\sigma}$ with probability $P_{acc}(\tilde{\sigma}) = \sum_{i \in F} \mu_i$. If $f(\tilde{\sigma}) = 1$, then we have $P_{acc}(\tilde{\sigma}) \ge 1/2 + \varepsilon$, else we have $P_{acc}(\tilde{\sigma}) \le 1/2 - \varepsilon$.

We assume without loss of generality that P reads inputs in the natural order $x_1, \dots, x_n$. We consider all inputs $\tilde{\sigma}_n, \dots, \tilde{\sigma}_1$, such that $\tilde{\sigma}_i = \tilde{\sigma}_i^0 \tilde{\sigma}_i^1$, where $\tilde{\sigma}_i^0 = \underbrace{0 \dots 0}_{n-i}$, $\tilde{\sigma}_i^1 = \underbrace{1 \dots 1}_{i}$.

For $i \in \{1, \dots, n\}$ we denote by μ^i a probability distribution after reading the part $\tilde{\sigma}_i^0$. That is, $\mu^i = \mu^0 A(0) \cdots A(0) = \mu^0 A^{n-i}(0)$. There are only ones in the $\tilde{\sigma}_i^1$, hence a computation after reading $\tilde{\sigma}_i^0$ can be described by a Markov chain. In this case μ^i is the initial probability distribution for a Markov process, $A(1)$ is the transition probability matrix. The states of this Markov chain are either ergodic or transient cf., e.g., [11]. An ergodic set of states is a set which a process cannot leave if it once entered. A transient set of states is a set which a process can leave, but cannot return if it once left. An ergodic state is an element of an ergodic set. A transient state is an element of transient set.

An arbitrary Markov chain C has at least one ergodic set. C can be a Markov chain without any transient set. If a Markov chain C has more than one ergodic set, then there is no interaction between these sets. Hence we have two or more unrelated Markov chains lumped together. Those chains can be studied separately. If a Markov chain consist of a single ergodic set, then the chain is called an ergodic chain. According to the usual classification, every ergodic chain is either regular or cyclic.

If an ergodic chain is regular, then a sufficiently high power of a state transition matrix A has only positive elements. Thus no matter where the process starts, after a sufficient number of steps it can be in any state. Moreover, there is a limiting vector of probabilities of being in the states of the chain which do not dependent on the initial state.

If a Markov chain is cyclic, then a chain has a period t and all its states are subdivided into t cyclic subsets ($t > 1$). For a given starting state a process moves through the cyclic subsets in a definite order, returning to the subset with the starting state after every t steps. It is known that after the sufficient time elapsed, the process can be in any state of a cyclic subset appropriate at the given moment. Hence for each of t cyclic subsets, the

t-th power of a state transition matrix A^t describes a regular Markov chain. Moreover, if an ergodic chain is a cyclic chain with a period t, it has at least t states.

From the assumption $q < p_n$ in the proof, we get that $t < p_n$ for every cyclic chain. We denote by D the least common multiply of all such t. Because p_n is prime, t is relatively prime to p_n, D is relatively prime to p_n, and so is any positive power D^m of D.

Let α^k be a probability distribution after reading the part $\tilde{\sigma}_k^1$ of $\tilde{\sigma}_k$. That is, $\alpha^k = \mu^k A^k(1)$. We can assume that there is a single accepting state, without loss of generality. Let α_{acc}^k be the probability of being in accepting state after reading the input $\tilde{\sigma}_k^1$. Since after every D steps a process can be in any state comprising an accepting state, D-th power of matrix A describes a regular Markov chain for that set. From the theory of Markov chains we have that there exists α_{acc} that $\lim_{k\to\infty} \alpha_{acc}^{kD} = \alpha_{acc}$. Hence for any $\varepsilon > 0$ it holds that

$$|\alpha_{acc}^{D^m} - \alpha_{acc}^{D^m p_n}| < 2\epsilon$$

for m large enough. As P $(1/2+\varepsilon)$-computes MOD_{p_n}, we have that $\alpha_{acc}^{D^m p_n} \geq 1/2 + \varepsilon$ and $\alpha_{acc}^{D^m} \leq 1/2 - \varepsilon$, a contradiction with the inequality above.

□

4 Lower Bounds

Below we present a general lower bound on the width of 1QBPs and compare it with the width of deterministic OBDDs computing the same function.

Theorem 3 *Let $\varepsilon \in (0, 1/2)$. Let f_n be a Boolean function which is $(1/2 + \varepsilon)$-computed (computed with a margin ε) by a 1QBP Q. Then it holds that*

$$width(Q) = \Omega\left(\frac{\log width(P)}{\log\log width(P)}\right)$$

where P is a deterministic OBDD of minimal width which computes f_n.

The next theorem presents a more precise lower bound for a particular margin ε of computation.

Theorem 4 *Let $\varepsilon \in (3/8, 1/2)$. Let f_n be a Boolean function which is $(1/2 + \varepsilon)$-computed (computed with a margin ε) by a 1QBP Q. Then it holds that*

$$width(Q) = \Omega\left(\frac{\log width(P)}{2\log(1+1/\tau)}\right)$$

where P is a deterministic OBDD of minimal width computing f_n and $\tau = \sqrt{1 + 2\varepsilon - 4\sqrt{1/2 - \varepsilon}}$.

Proofs of the above theorems are presented in the section below.

4.1 Proofs of Theorems 3 and 4

Proofs of Theorems 3, and 4 use a similar idea. We construct a deterministic OBDD P that computes the same function f_n and

$$width(P) \leq \left(1 + \frac{2}{\theta}\right)^{2\,width(Q)}. \qquad (1)$$

Proofs of Theorems 3, 4 differ only in an estimation of a parameter $\theta > 0$ depending on ε.

4.1.1 A Deterministic OBDD-Presentation of a 1QBP

Let $d = width(Q)$. Let $\pi = \{i_1, i_2, \ldots, i_n\}$ be an ordering of testing variables of Q. From now on we assume that the input sequences $\sigma \in \{0,1\}^n$ are ordered by an order π determined by Q. We define a deterministic OBDD *LQ* based on Q as follows. *LQ* use the ordering π of testing variables represented by the following labeled complete binary tree.

- The initial node of *LQ* is marked by an initial configuration $|\psi_0\rangle$ of Q. Two outgoing vertices of the initial node are marked by $x_{i_1} = 1$ and $x_{i_1} = 0$.
- Two nodes of *LQ* on the level 1 are marked by the configurations $|\psi_1(0)\rangle$ and $|\psi_1(1)\rangle$ of Q where $|\psi_1(\sigma_1)\rangle$ is the configuration after the first step of computation after reading $x_{i_1} = \sigma_1$ for $\sigma_1 \in \{0,1\}$.

  A vertex $x_{i_1} = \sigma_1$ leads from the node $|\psi_0\rangle$ to the node $|\psi_1(\sigma_1)\rangle$ iff $|\psi_1(\sigma_1)\rangle = U_1(\sigma_1)|\psi_0\rangle$.
- Consider a level j of *LQ*. 2^j nodes of *LQ* of the level j are marked by the configurations $\{|\psi_j(\sigma_1 \ldots \sigma_j)\rangle \in \Psi : \sigma_1 \ldots \sigma_j \in \{0,1\}^j\}$ where $|\psi_j(\sigma_1 \ldots \sigma_j)\rangle$ is a configuration of Q after reading the first part $\sigma_1 \ldots \sigma_j$ of the input $\sigma \in \{0,1\}^n$.

A vertex (marked by $x_{i_{j+1}} = \sigma_{j+1}$) from the node $|\psi_j(\sigma_1 \dots \sigma_j)\rangle$ leads to the node $|\psi_{j+1}(\sigma_1 \dots \sigma_j \sigma_{j+1})\rangle$ iff $|\psi_{j+1}(\sigma_1 \dots \sigma_j \sigma_{j+1})\rangle = U_{j+1}(\sigma_{j+1})|\psi_j(\sigma_1 \dots \sigma_j)\rangle$.

- Consider the last level n of LQ. We mark 2^n nodes of LQ on the level n by the configurations $|\psi_n(\sigma_1 \dots \sigma_n)\rangle \in \Psi : (\sigma_1 \dots \sigma_n) \in \{0,1\}^n$ and in addition we mark them by 0 and 1 as follows. We mark node $|\psi_n(\sigma_1 \dots \sigma_n)\rangle$ by 1 if for configuration $|\psi_n(\sigma_1 \dots \sigma_n)\rangle$ it holds that $p_{accept}(\sigma_1 \dots \sigma_n) \geq 1/2 + \varepsilon$. We mark a node $|\psi_n(\sigma_1 \dots \sigma_n)\rangle$ by 0 if for configuration $|\psi_n(\sigma_1 \dots \sigma_n)\rangle$ it holds that $p_{accept}(\sigma_1 \dots \sigma_n) \leq 1/2 - \varepsilon$.

Property 4 *A deterministic OBDD LQ computes the same Boolean function f_n as Q.*

Proof: Evident and follows from the construction of LQ. □

4.1.2 A Metric Automaton Characterization of LQ

We view now an OBDD LQ with an ordering π of testing variables as the following metric time-variant automaton that reads its input sequences $\sigma \in \{0,1\}^n$ in an order π:

$$LQ = \langle \{0,1\}, \Psi, \{\delta_j\}_{j=1}^n, |\psi_0\rangle, \mathcal{F}_\varepsilon \rangle$$

where $\{0,1\}$ is the input alphabet, $\Psi = \{|\psi\rangle\}$ is a set of states (set of all possible configurations of Q during its computations on inputs from $\{0,1\}^n$). That is, $\Psi = \cup_{j=0}^n \Psi_j$ where Ψ_j is a set of states of LQ on the level j. An automaton transition function $\delta_j : \Psi_{j-1} \times \{0,1\} \to \Psi_j$ determines transitions in the step j, $1 \leq j \leq n$, (δ_j is defined according to the transitions of LQ in the level $j-1$). Finally $|\psi_0\rangle$ is the initial state and $\mathcal{F}_\varepsilon = \{|\psi\rangle \in \Psi_n : ||M|\psi\rangle||^2 \geq 1/2 + \varepsilon\}$ is the accepting set of states of LQ.

For $j \in \{1, \dots, n\}$, we denote by $\Delta_j : \Psi_{j-1} \times \{0,1\}^{n-j+1} \to \Psi_n$ the automaton transitive closure of the sequence $\delta_j, \dots, \delta_n$ of the transition functions. That is,

$$\Delta_j(|\psi\rangle, \sigma_j \dots \sigma_n) = \delta_n(\dots(\delta_j(|\psi\rangle, \sigma_j), \dots, \sigma_n).$$

Lemma 4 *Let f_n be a Boolean function $(1/2+\varepsilon)$-computed by a LQ. Let $\theta > 0$, and for arbitrary $|\psi\rangle \in \mathcal{F}_\varepsilon$ and arbitrary $|\psi\rangle' \in \Psi_n \backslash \mathcal{F}_\varepsilon$ it holds that*

$$||\,|\psi\rangle - |\psi'\rangle\,|| \geq \theta.$$

Then, there exists a deterministic OBDD B which computes f_n and

$$width(B) \leq \left(1 + \frac{2}{\theta}\right)^{2d}.$$

Proof: We recall first some known notions concerning metric spaces (see [3]). A Hilbert space $\mathcal{H}_d$ is a metric space with a metric defined by the norm $||\cdot||$. The points μ, μ' from $\mathcal{H}_d$ are connected through a θ-chain if there exists a finite set of points $\mu_1, \mu_2, \ldots, \mu_m$ from $\mathcal{H}_d$ such that $\mu_1 = \mu, \mu_m = \mu'$ and $||\mu_i - \mu_{i+1}|| \leq \theta$ for $i \in \{1, \ldots, m-1\}$. A subset C of $\mathcal{H}_d$ is called a θ-component if arbitrary two points $\mu, \mu' \in \mathcal{C}$ are connected through a θ-chain. It is known [3] that if $\mathcal{D}$ is a finite diameter subset of a subspace of $\mathcal{H}_d$ (a diameter of $\mathcal{D}$ is defined as $\sup_{\mu,\mu' \in \mathcal{D}}\{||\mu - \mu'||\}$ then for $\theta > 0$ $\mathcal{D}$ is partitioned to a finite number t of its θ-components.

A set Ψ of states of LQ belongs to the sphere of radius 1 which has center $(0, 0, \ldots, 0)$ in $\mathcal{H}_d$ because for all $|\psi\rangle \in \Psi$ it holds that $||\,|\psi\rangle\,|| = 1$. For each $j \in \{0, \ldots, n\}$, denote by $[\Psi_j] = \{C_1, \ldots, C_{t_j}\}$ the set of θ-components of $\Psi_j \subset \mathcal{H}_d$.

From the condition of our Lemma it follows that a subset $\mathcal{F}_\varepsilon$ of Ψ_n is a union of some θ-components of Ψ_n. The transition functions δ_j, $1 \leq j \leq n$, preserve the distance. That is, for arbitrary $|\psi\rangle$ and $|\xi\rangle$ from Ψ_j and arbitrary $\gamma \in \{0, 1\}$, it holds that

$$||\,|\psi\rangle - |\xi\rangle\,|| = ||\delta_j(|\psi\rangle, \gamma) - \delta_j(|\xi\rangle, \gamma)||. \tag{2}$$

From (2) we have that for $C \in [\Psi_j]$ and for $\gamma \in \{0, 1\}$ there exists $C' \in [\Psi_{j+1}]$ such that $\delta_j(C, \gamma) = C'$. Here $\delta_j(C, \gamma)$ is defined as $\delta_j(C, \gamma) = \cup_{|\psi\rangle \in C}\delta_j(|\psi\rangle, \gamma)$.

Now we describe a deterministic OBDD B in terms of a time-variant finite automaton that computes f_n.

$$B = \langle \{0, 1\},\, [\Psi],\, \{\delta_j\}_{j=1}^n,\, C_0,\, F\rangle$$

where $[\Psi] = \cup_{j=0}^n [\Psi_j]$ is a set of states of B ($[\Psi_j]$ is a set of states on the step j of a computation of B);
$\delta_j : [\Psi_{j-1}] \times \{0, 1\} \to [\Psi_j]$ is a transition function of B in the step j;
an initial state $C_0 = \{|\psi_0\rangle\}$ is a one-element θ-component of Ψ_0; we define F by $F = \{C_i \in [\Psi_n] : C_i \subseteq \mathcal{F}_\varepsilon\}$.

From our construction we have that B and LQ compute the same function f_n. The width of B is $t = \max\{t_0, \ldots, t_n\}$.

Let $t = t_j$ We estimate a number t of θ-components (number of states of B) of Ψ_j as follows. For each θ-component C, we select one point $|\psi\rangle \in C$.

If we draw a sphere of the radius $\theta/2$ with the center $|\psi\rangle \in C$ then all such spheres do not intersect pairwise. All these spheres (t many) are in a larger sphere of radius $1 + \theta/2$ which has center $(0, 0, \ldots, 0)$. The volume of a sphere of radius r in $\mathcal{H}_d$ is cr^{2d}, where the constant c depends on a metric of $\mathcal{H}_d$. Note that for estimating the volume of the sphere we should take into account that $\mathcal{H}_d$ is a d-dimensional *complex* space and each complex point is a 2-dimensional point. So it holds that

$$width(B) \leq \frac{c\,(1+\theta/2)^{2d}}{c\,(\theta/2)^{2d}} = \left(1 + \frac{2}{\theta}\right)^{2d}.$$

$\square$

Below we formulate a technical lemma that estimates a number of components of Ψ for different ε.

Lemma 5 *Let an LQ $(1/2+\varepsilon)$-computes a function f_n. Then for arbitrary $|\psi\rangle \in \mathcal{F}_\varepsilon$ and arbitrary $|\psi'\rangle \notin \mathcal{F}_\varepsilon$ it holds that*

1. $||\,|\psi\rangle - |\psi'\rangle\,|| \geq \theta_1 = \varepsilon/\sqrt{d}$

   and

2. $||\,|\psi\rangle - |\psi'\rangle\,|| \geq \theta_2 = \sqrt{1 + 2\varepsilon - 4\sqrt{1/2 - \varepsilon}}.$

Proof: For the sake of simplification we denote a configuration $|\psi\rangle = z_1|1\rangle + \cdots + z_d|d\rangle$ just by $\psi = (z_1, \ldots, z_d)$. Let $\psi = (z_1, \ldots, z_d)$ and $\psi' = (z'_1, \ldots z'_d)$. Consider a norm $||.||_1$ defined as $||\psi||_1 = \sum_{i=1}^{d} |z_i|$.

1. From the definition of LQ it holds that

$$2\varepsilon \leq \sum_{s_i \in F} (|z_i|^2 - |z'_i|^2) = \sum_{s_i \in F} (|z_i| - |z'_i|)(|z_i| + |z'_i|) \leq$$

$$\leq 2 \sum_{s_i \in F} (|z_i| - |z'_i|) \leq 2 \sum_{s_i \in F} |z_i - z'_i| \leq 2||\psi - \psi'||_1$$

Using an inequality

$$a_1b_1 + a_2b_2 + \ldots + a_db_d \leq \sqrt{a_1^2 + a_2^2 + \ldots + a_d^2}\sqrt{b_1^2 + b_2^2 + \ldots + b_d^2}, \qquad (3)$$

for $b_1 = b_2 = \ldots = b_d = 1$ we get that $||\psi||_1 \leq \sqrt{d}||\psi||$. Therefore,

$$2\varepsilon \leq 2||\psi - \psi'||_1 \leq 2\sqrt{d}||\psi - \psi'||$$

Finally, we have

$$||\psi - \psi'|| \geq \varepsilon/\sqrt{d}.$$

2. Consider now the next variant of a lower bound for $||\psi - \psi'||$.

$$\begin{aligned}
||\psi - \psi'|| &= \sqrt{\sum_{i=1}^{d} |z_i - z_i'|^2} \geq \sqrt{\sum_{i=1}^{d} (|z_i| - |z_i'|)^2} = \\
&= \sqrt{\sum_{i=1}^{d} |z_i|^2 + \sum_{i=1}^{d} |z_i'|^2 - 2\sum_{i=1}^{d} |z_i||z_i'|} \geq \\
&\geq \sqrt{\sum_{s_i \in F} |z_i|^2 + \sum_{s_i \notin F} |z_i'|^2 - 2\sum_{s_i \in F} |z_i||z_i'| - 2\sum_{s_i \notin F} |z_i||z_i'|}.
\end{aligned}$$

From the definition of an LQ we have that $\sum_{s_i \in F} |z_i|^2 \geq 1/2 + \varepsilon$, $\sum_{s_i \notin F} |z_i'|^2 \geq 1/2 + \varepsilon$. Now, from the above we get that

$$||\psi - \psi'|| \geq \sqrt{1/2 + \varepsilon + 1/2 + \varepsilon - 2\sum_{s_i \in F} |z_i||z_i'| - 2\sum_{s_i \notin F} |z_i||z_i'|}.$$

Using inequality (3) we get

$$||\psi - \psi'|| \geq \sqrt{1 + 2\varepsilon - 2\sqrt{\sum_{s_i \in F} |z_i|^2}\sqrt{\sum_{s_i \in F} |z_i'|^2} - 2\sqrt{\sum_{s_i \notin F} |z_i|^2}\sqrt{\sum_{s_i \notin F} |z_i'|^2}}.$$

Using the property $\sum_{s_i \notin F} |z_i|^2 \leq 1/2 - \varepsilon$, $\sum_{s_i \in F} |z_i'|^2 \leq 1/2 - \varepsilon$, $\sum_{s_i} |z_i|^2 \leq 1$, and $\sum_{s_i} |z_i'|^2 \leq 1$, we finally get that

$$||\psi - \psi'|| \geq \sqrt{1 + 2\varepsilon - 4\sqrt{1/2 - \varepsilon}} = \theta_2.$$

$\square$

Note that the lower bound above for $||\psi - \psi'||$ is nontrivial (positive) if $\varepsilon \in (\alpha, 1/2)$ where α is about 3/8. For $\varepsilon \in (0, \alpha]$ it holds that $1 + 2\varepsilon - 4\sqrt{1/2 - \varepsilon} \leq 0$. In that case the lower bound $||\psi - \psi'|| \geq \varepsilon/\sqrt{d}$ is more precise.

Now we turn to the formal estimations of the lower bounds of Theorems 3 and 4.

Proof of Theorem 3: From Lemma 4 and Lemma 5, it follows that

$$t \leq \left(1 + \frac{2\sqrt{d}}{\varepsilon}\right)^{2d}$$

or $\log t = O(d \log d)$. From that we get that

$$d = \Omega\left(\frac{\log t}{\log\log t}\right).$$

Proof of Theorem 4: From Lemma 4 and Lemma 5, it follows that

$$t \leq \left(1 + \frac{2}{\theta_2}\right)^{2d}$$

or $2d \geq \log t / \log(1 + 2/\theta_2)$. From this we have that

$$d \geq \frac{\log t}{2\log(1 + 1/\theta_2)}.$$

□

References

[1] F. Ablayev and M. Karpinski, On the Power of Randomized Branching programs, *Electronic Colloquium in Computational Complexity*, ECCC TR98-004, (1998), available at `http://www.eccc.uni-trier.de/eccc/` , also appeared in Proc. 28th ICALP (1996), LNCS Vol. 1099, Springer, 1996, 348-356.

[2] F. Ablayev and M. Karpinski, A Lower Bound for Integer Multiplication on Randomized Read-Once Branching Programs, ECCC TR98-011 (1998), available at `http://www.eccc.uni-trier.de/eccc/`.

[3] P. Alexandrov, Introduction to Set Theory and General Topology, Berlin, 1984.

[4] A. Ambainis and R. Freivalds, 1-Way Quantum Finite Automata: Strengths, Weaknesses and Generalization, In *Proceeding of the 39th*

IEEE Conference on Foundation of Computer Science, 1998, 332-342. See also quant-ph/9802062 v3.

[5] Bernstein and Vazirany, Quantum Complexity Theory, SIAM J. Comput, Vol. 26, No. 5, (1997), 1411-1473. Preliminary version of this paper appeared in Proc. 25th STOC, (1993), 11-20.

[6] D. Barrington, Bounded-width Polynomial-Size Branching Programs Recognize Exactly Those Languages in NC^1, *Journal of Comp. and System Sci.* 38, (1989), 150-164.

[7] A.Brodsky and N.Pippenger. Characterization of 1-Way Quantum Finite Automata. *quant-ph/9903014*, available at `http://xxx.lanl.gov/archive/quant-ph`. See also its Russian mirror: `http://xxx.itep.ru`.

[8] F. Green, S. Homer, and Pollet, On the Complexity of Quantum ACC, *in Proc. of the 15th IEEE Conf. on Computational Complexity* 2000.

[9] R. Feynman, Simulation Physics with Computers, *International Journal of Theoretical Physics*, (1982), 21, 467.

[10] R. Freivalds, Quantum Finite Automata, *Manuscript 2000*, personal communication.

[11] J. G. Kemeny, J. L. Snell, Finite Markov Chains, Van Nostrand Company, INC, 1960.

[12] A.Kondacs and J.Watrous. On the Power of Quantum Finite State Automata. In *proceedings of the 38th Annual Symposium on Foundations of Computer Science*, 1997, 66-75.

[13] C. Moore and M. Nilson, Some Notes on Parallel Quantum Computing, *quant ph/9804034*.

[14] M.Nakanishi, K.Hamaguchi, and T.Kashiwabara, Ordered quantum branching programs are more powerful than ordered probabilistic branching programs under a bounded-width restriction, *in the proceedings of the 6th Annual International Conference on Computing and Combinatorics, COCOON'2000, Lecture Notes in Computer Science, Springer-Verlag*, v. 1858, (2000), 467-476.

[15] P. Shor, Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer, *SIAM J. on Computing*, 26(5), (1997), 1484-1509.

[16] B.C. Travaglione, M.A. Nielsen, H.M. Wiseman, and A. Ambainis, ROM-based computation: quantum versus classical, *arXiv: quant-ph/0109016 4 Sep 2001*.

[17] I. Wegener, *Branching Programs and Binary Decision Diagrams.* Siam Monographs on Discrete Mathematics and Applications. 2000.